

**Method and Apparatus for Learning Probabilistic Relational
Models Having Attribute and Link Uncertainty and for
Performing Selectivity Estimation Using Probabilistic
Relational Models**

5

ABSTRACT

Sub
AI

10 The invention comprises a method and apparatus for learning probabilistic
models (PRM's) with attribute uncertainty. A PRM with attribute uncertainty
defines a probability distribution over instantiations of a database. A learned
PRM is useful for discovering interesting patterns and dependencies in the
data. Unlike many existing techniques, the process is data-driven rather than
hypothesis driven. This makes the technique particularly well-suited for
exploratory data analysis. In addition, the invention comprises a method and
15 apparatus for handling link uncertainty in PRM's. Link uncertainty is
uncertainty over which entities are related in our domain. The invention
comprises of two mechanisms for modeling link uncertainty: reference
uncertainty and existence uncertainty. The invention includes learning
algorithms for each form of link uncertainty. The third component of the
20 invention is a technique for performing database selectivity estimation using
probabilistic relational models. The invention provides a unified framework for
the estimation of query result size for a broad class of queries involving both
select and join operations. A single learned model can be used to efficiently
estimate query result sizes for a wide collection of potential queries across
25 multiple tables.